

文章编号: 1003-0077(2018)08-0128-07

基于协同过滤 Attention 机制的情感分析模型

赵冬梅^{1,2}, 李 雅², 陶建华², 顾明亮¹

(1. 江苏师范大学 物理与电子工程学院, 江苏 徐州 221116;
2. 中国科学院 自动化研究所 模式识别国家重点实验室, 北京 100101)

摘 要: 该文主要研究在评论性数据中用户个性及产品信息对数据情感类别的影响。在影响数据情感类型的众多因素中, 该文认为评价的主体即用户以及被评价的对象等信息对评论数据的情感至关重要。该文提出一种基于协同过滤 Attention 机制的情感分析方法(LSTM-CFA), 使用协同过滤(CF)算法计算出用户兴趣分布矩阵, 再将矩阵利用 SVD 分解后加入层次 LSTM 模型, 作为模型注意力机制提取文档特征、实现情感分类。实验表明 LSTM-CFA 方法能够高效提取用户个性与产品属性信息, 显著提升了情感分类的准确率。

关键词: 情感分析; 协同过滤; LSTM; 注意力机制; SVD

中图分类号: TP391

文献标识码: A

Sentiment Analysis Based on Collaborative Filter Attention Mechanism

ZHAO Dongmei^{1,2}, LI Ya², TAO Jianhua², GU Mingliang¹

(1. School of Physics Electronic Engineering, Jiangsu Normal University, Xuzhou, Jiangsu 221116, China;
2. National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100101, China)

Abstract: This paper investigates the influence of user's personality and product information on data emotion category in review data. Among the many factors that affect the emotional data type, the subject of the evaluation, that is, the user and the object of the evaluation, are emotionally important to the commentary data. In this paper, an emotional analysis model (LSTM-CFA) based on cooperative filtering attention mechanism is proposed. The user interest distribution matrix is calculated by using the collaborative filtering (CF) algorithm. After the matrix is decomposed with SVD, the matrix is added to the hierarchical LSTM model as an attention mechanism in order to achieve emotion classification. Experiments show that the LSTM-CFA model can extract the information of user's personality and product attribute efficiently, to improve the accuracy of emotion classification.

Key words: sentiment analysis; collaborative filtering; LSTM; attention mechanism; SVD

0 引言

情感分析旨在从文本中挖掘出用户对某一热点问题或者产品性能的观点信息。在实际应用中不仅可以帮助电商企业评估产品的市场价值, 也可以帮助政府部门做好舆情监控等任务。因此, 文本情感分析越来越受到关注, 并成为自然语言处理领域的重要课题之一^[1]。

在传统的情感分析中, 研究者们往往只关注评

论内容的重要性。根据绝大多数推荐网站上产品评论信息中提供的评价用户信息及其喜好信息, 可以发现同一用户评论数据的情感极性比来自不同用户的情感极性更倾向于一致。因此, 将用户及被评价的产品信息融入情感分类的任务中至关重要。

目前已有研究者考虑在模型中加入用户及产品信息。例如, 将用户、产品以及评论数据特征基于词袋模型输入不同的分类器中进行探究^[2]。也有使用概率模型获取用户的兴趣分布和文本内容分布^[3]。而随着深度学习在计算机视觉、语音识别、自然语言

收稿日期: 2017-10-27 定稿日期: 2017-11-17

基金项目: 国家自然科学基金(61425017, 61403386, 61773379); “863”高技术研究计划(2015AA016305); 国家社会科学基金(13&ZD189)

处理领域的成功应用,基于深度学习的模型越来越成为情感分析的主流方法。通过神经网络可以提取出文本中更丰富的语义信息以及包含的用户、产品信息。但目前大多数模型都是将用户和产品信息加入文本中一起输入模型训练,这样不仅使评论内容失去原本的语义,且导致用户个性等信息也没能被提取出来。

由于成长环境、教育背景等差异,不同用户的性格习惯存在一定区别。评价产品时,用户会对喜欢的物品打高分,对不喜欢的物品打低分,但也存在用户对喜欢或不喜欢的物品都选择低分或者高分,这就凸显了用户的个性信息。本文提出 LSTM-CFA 模型,通过使用推荐系统中协同过滤算法从用户和产品信息中提取出用户兴趣分布信息,并将此信息作为模型的注意力机制。在使用层次 LSTM 模型的过程中,指导词语级别和句子级别的特征表示,以提高文档级别的情感分类任务。

本文的创新点主要有以下两个方面:

(1) 使用协同过滤算法,充分挖掘用户的个性以及产品信息,将其作为模型的注意力机制。

(2) 使用奇异值分解(SVD)对稀疏矩阵进行优化分解,得到用户个性矩阵和产品信息矩阵。

1 相关工作介绍

情感分析任务作为自然语言处理领域的重要课题之一,旨在从文本中挖掘出主题或者用户观点等情感信息。划分出的情感类别可以是积极和消极,也可以像影评数据集中那样使用 0—10 表示级别。

近几年,由于神经网络模型有较好的文本表示能力,开始被引用到自然语言处理领域解决实际问题。Hinton 最早提出将非结构化的文本映射到低维实值空间中^[4],演变成现在的词向量。词向量可以从文本中提取有效的语义信息并计算出词语之间的相似性。Mikolov 提出 Word2Vec 工具将获取词向量方法变得更加方便有效^[5],随后又提出了基于文档的 Paragraph 工具,给文档表示提供了很大的帮助。

在情感分类任务中,Socher 等人使用一系列的递归神经网络学习文本表示方法,包括递归自编码模型(recursive autoencoders, RAE)^[6]、矩阵向量递归神经网络(matrix-vector recursive neural network, MVRNN)^[7]、递归神经张量网络(recursive neural tensor network, RNTN)^[8],在斯坦福语料

(Stanford sentiment treebank, SST)上获得 5.4% 提高。Kim 参考图像领域的卷积神经网络模型(convolutional neural network, CNN)修改并加入滑动窗口提取不同的文本特征,用来解决文本分类以及情感分类问题,取得很好的效果^[9]。Kalchbrenner 等人提出动态的 CNN 模型,使用动态的池化层处理不同句长的句子^[10]。另外,循环神经网络(recursive neural network, RNN)在时序数据上的不俗表现使其在自然语言处理领域广受欢迎。Tai 等人建立了树结构的 LSTM 模型用来处理情感分类任务^[11]。Tang 等人构建层次的 LSTM 模型来表示不同级别的特征信息,在文档分类任务中效果较好^[12]。注意力模型可以模拟人类注意力运行机制在图像领域取得不俗的表现,在自然语言处理领域中可以通过注意力机制从词语或句子中提取出更重要的信息作为特征^[13]。

2 模型及算法介绍

对比已有的基于用户和产品的情感分析模型^[14],本文提出了基于协同过滤注意力机制的情感分析模型(long short time memory cooperative filter attention, LSTM-CFA)。使用推荐系统中的协同过滤算法(cooperative filter, CF)以获得用户兴趣分布矩阵,再将此矩阵作为 LSTM 模型的注意力机制(attention),模型结构如图 1 所示。实验主要流程为:首先将训练好的词向量作为第一层 LSTM 模型输入,将得到的隐藏层结合注意力机制表示成句子向量,再把句子向量通过第二层 LSTM 模型得到句子隐藏层,然后再次结合注意力机制得到文档表示特征,通过全连层后利用 softmax 层完成情感分析工作。

2.1 基于层次 LSTM 特征表示模型

循环神经网络可以通过链式神经网络结构传播历史信息。在处理时序数据时,它每次能看到当前的输入 x_t 以及上一时刻输出的隐藏状态 h_{t-1} 。然而在训练 RNN 时,发现反向传播算法会导致在长距离传输过程中梯度弥散,当前信息无法传递到时间 t 以外的节点处,降低对历史信息的感知力^[15]。为了解决这个问题,Hochreiter 等提出长短期记忆模型(LSTM)^[16],并在自然语言处理领域有着很好的应用效果。

因此,本文使用层次 LSTM 分别表示句子级别

特征和文档级别特征。首先使用 Word2Vec 预训练得到词向量表示形式,将句中词语嵌入低维语义空间。每一步中,使用词向量 x_t 作为输入,此时的隐藏状态 h_t 和当前记忆门 c_t 可以由之前的门状态 c_{t-1} 和隐藏状态 h_{t-1} 更新得到。具体过程如式(1)~式(6)所示。

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (1)$$

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

$$q_t = \tanh(W_q \cdot [h_{t-1}, x_t] + b_q) \quad (3)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (4)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot q_t \quad (5)$$

$$h_t = o_t \odot \tanh(c_t) \quad (6)$$

其中, σ 表示 sigmoid 函数。 \tanh 表示双曲正切函数, \odot 表示矩阵元素点乘操作, f_t 是控制丢弃旧信息记忆门信息的程度函数, i_t 是控制从当前记忆门中存储多少新信息的函数, o_t 是根据当前门状态 c_t 控制输出信息。 x_t 可以表示某一篇文章中单词,所以在此也可以写成 x_t^i , 得到的隐藏层便可以写成 h_t^i 。所有的词语隐藏状态 h_t 可以通过池化层得到句子特征,在本文中使用注意力机制监督得到句子级别特征表示,具体结合方式在 2.2.1 节中给出。

同理,句子级别特征通过第二层 LSTM 模型后便可以得到文档的特征表示。

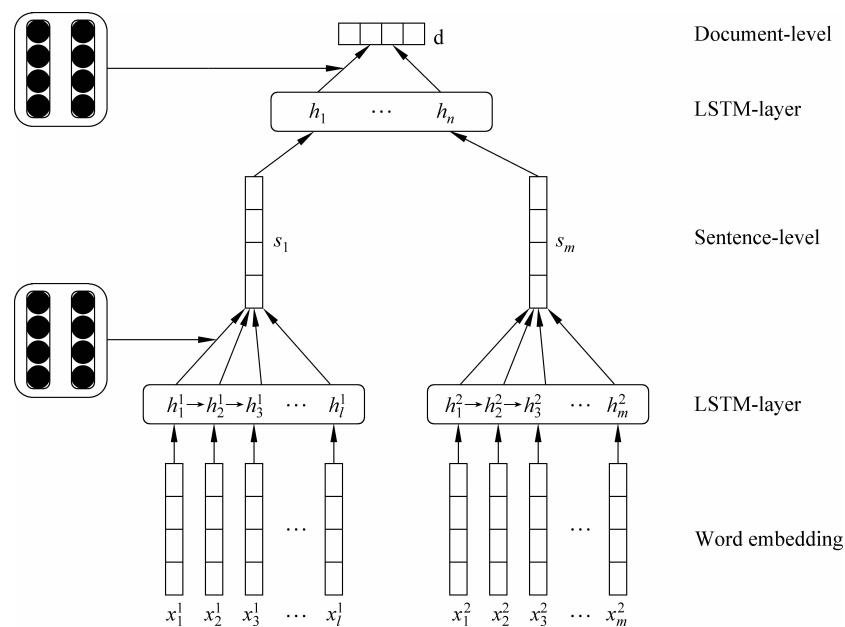


图1 基于协同过滤的情感分析模型框架

2.2 基于协同过滤的注意力机制模型

协同过滤算法已被成功运用于许多推荐系统中^[17]。主要分为,基于用户相似性的协同过滤算法和基于物品相似性的协同过滤算法。本文使用基于物品相似度的方法,通过用户对已用的物品评分来计算用户对与该物品相似的物品评分,以预测用户偏好分布并推荐物品。

2.2.1 用户兴趣分布矩阵计算

本文希望通过协同过滤算法获得用户兴趣分布矩阵作为情感分析模型的注意力机制,文中矩阵计算步骤如下。

1) 构建物品之间的共现矩阵

共现矩阵是表示某一用户喜欢或评分的物品列表集合。如用户 A 评分过的物品有 a、b、c、d,那么

对于物品 a 而言, b、c、d 均与 a 共同存在过,即在对应的矩阵中标 1。同理,可以得到物品数量大小的方阵,其中数值表示两物品是否在同一用户下存在过。对于得到的方阵数值只有 0 或者 1 表示,为了计算方便,这里使用余弦相似度对共现矩阵做归一化处理,如式(7)所示。

$$\cos Sim(a, b) = W_b^a = \frac{\langle a, b \rangle}{||a|| ||b||} \quad (7)$$

其中 W_b^a 表示物品 a 与 b 之间的相似度。

2) 建立用户对物品的评分矩阵

将用户对物品的评分构建成矩阵。矩阵行表示用户,列表示物品,数值是某一用户对某物品的打分情况。若没有打分则记为 0,打分则记为相应分数,矩阵形式如表 1 所示,其中 ui 列表示用户($i=1 \dots 4$),pj 行表示产品($j=1 \dots 5$)。

表 1 评分矩阵形式

	p1	p2	p3	p4	p5
u1	3	0	5	0	0
u2	0	9	8	0	7
u3	5	0	5	3	6
u4	0	4	3	0	2

3) 计算得到用户兴趣分布矩阵

根据物品相似度以及打分情况通过式(8)便可以计算出用户对没有打分的物品的喜好程度。

$$r_{uj} = \frac{\sum W_j^i \cdot r_{ui}}{\sum |W_j^i|} \quad (8)$$

其中, r_{ui} 表示用户对物品 i 的打分。

4) 用户兴趣分布矩阵分解优化

根据以上三个步骤便可以计算出用户兴趣分布矩阵,该矩阵维度大小为用户数量乘以产品数量。当数据较多的情况下高维矩阵不仅会影响计算速度,也会存在稀疏可能。本文使用奇异值分解(SVD)分解法对高维矩阵进行分解处理^[18],得到用户兴趣分布矩阵和产品属性分布矩阵,如式(9)所示。

$$A = UV^T \quad (9)$$

其中, A 表示待分解矩阵,在这里指用户兴趣分布矩阵, U 与 V^T 是分解后得到的矩阵。在本文中,表示用户个性矩阵和产品属性矩阵。

2.2.2 注意力机制模型

注意力机制(attention)是松散的基于人脑注意力的一种机制,在文本中通过自动加权的方法能够有效的捕捉到文本中重要的信息。本文中使用用户兴趣分布矩阵作为注意力模型从隐藏状态中提取对句子贡献较大的词语,用这些词表征句子的特征。具体计算如式(10)~式(12)所示。

$$u_{ij} = \tanh(W_h h_j^i + W_u u + W_p p + b) \quad (10)$$

$$\alpha_{ij} = \frac{\exp(u_{ij}^T u_w)}{\sum_i \exp(u_{ij}^T u_w)} \quad (11)$$

$$s_i = \sum_j \alpha_{ij} h_{ij} \quad (12)$$

其中, W_h, W_u, W_p 是权重矩阵, b 表示偏置值, u, p 表示用户物品分布矩阵,可以通过奇异值分解用户兴趣分布矩阵得到。 u_{ij} 是打分函数,计算词语的重要程度。通过 softmax 函数计算 u_w 向量的权重值 α_{ij} 。最后将 α_{ij} 与对应的隐状态加权求和得到句子特征。

同理,句子在不同用户不同物品间表达的意义

也会有所不同,对于上述得到的句子特征后采用同样的权重计算,即可以得到文档表示特征 d ,如式(13)所示。

$$d = \sum_i \beta_i h_i \quad (13)$$

其中, β_i 表示句子中句子特征中第 i 个隐状态 h_i 的权重值,计算方式参考式(11)。

2.3 文档情感识别

本文使用 LSTM-CFA 模型对文档进行情感分析。使用预训练得到的词向量作为模型的输入数据,使用协同过滤得到用户兴趣分布矩阵作为模型的注意力机制。在层次 LSTM 模型中与隐藏层状态结合,提取重要信息组成文档特征。

使用交叉熵损失函数作为优化目标函数,使用反向传播算法计算并迭代更新模型参数,如式(14)所示。

$$Loss = - \sum_{d \in D} \sum_{c=1} p_c^e(d) \cdot \log(p_c(d)) \quad (14)$$

其中, D 表示训练集, c 表示情感类别, p_c^e 表示情感为 c 的 0,1 分布。当文本情感类别是 c 时, p_c^e 取值 1,否则 p_c^e 取值 0。 $p_c(d)$ 表示文档 d 情感类别是 c 时的概率大小。

3 实验

实验在三个带有用户和产品信息的情感数据集上对提出方法的有效性进行验证。其中,实验数据集的预处理部分主要将数据集按 8:1:1 比例分为训练集、验证集和测试集。

已有研究表明领域内语料可以让词向量拥有更多领域内的语义^[19],因此本文使用 Word2Vec 工具在实验数据上进行预训练词向量。实验环境具体配置数据如表 2 所示。

表 2 实验环境配置

实验环境	配置数值
预训练工具	Word2Vec
语言	Python2.7
框架	Theano 0.9.0
词向量维度	200 维

3.1 数据集介绍

数据集来自: IMDB, Yelp2013, Yelp2014^[20]。

其中 IMDB 数据集是用户对电影的评价内容,类别是类似豆瓣网站的星级评分制度,由最差 1 分至最好 10 分;后两者是用户对商家产品的评价信息。类似星级评价共有五个等级,从最差 1 分至最好 5 分制。数据集具体信息如表 3 所示。

表 3 数据集统计信息

数据集	类别	用户数	产品数	评论数据
IMDB	10	1 310	1 635	84 919
Yelp2013	5	1 631	1 633	78 966
Yelp2014	5	4 818	4 194	231 163

3.2 评价指标

本文使用精准率(accuracy)衡量分类的整体效果,用平方根误差(RMSE)衡量预测值和真值之间的离散程度。计算如式(15)~式(16)所示。

$$Accuracy = \frac{T}{N} \quad (15)$$

$$RMSE = \sqrt{\frac{\sum_i (gy - py)^2}{N}} \quad (16)$$

其中, T 表示预测正确类别的数量, N 表示文档总数。 gy 表示预测值, py 表示真实值。

3.3 实验分析

实验中的对比模型如下:

(1) Majority: 把训练集中用户使用主要的情感类别作为测试集数据的情感类别。

(2) Trigram: 使用 unigrams、bigrams 以及 trigrams 作为特征,输入到分类器 SVM 中。

(3) TextFeature: 人工设计文本特征后使用 SVM 分类。

(4) AvgWordvec+SVM: 使用 Word2Vec 学习词向量平均化后作为分类器 SVM 输入数据^[21]。

(5) SSWE(sentiment specific word embeddings)+SVM: 使用 Word2Vec 学习特定的情感词向量后经过池化层的最大化平均化后输入 SVM 中。

(6) Paragraph Vector: 使用 PVDN 解决文档分类问题^[20]。

(7) RNTN+Recurrent: 使用 RNTN 表示句子特征然后输入到 RNN 中分类^[8]。

(8) LSTM: 使用实验中 LSTM 模型,没有加入注意力机制。

(9) UPF(user product feature): 从训练数据

集中提取用户信息特征和产品特征,然后联合模型(2)~(3)进行分类^[14]。

(10) UPNN(user product neural network): 使用用户和产品信息输入 CNN 模型中分类^[22]。

(11) UPDMN(user product deep memory network): 使用深度记忆网络融合用户和产品信息^[23]。

(12) LSTM+CBA (cognition based attention): 基于注意力的认知模型结合 LSTM 模型分类^[24]。

(13) LSTM-CFA(Cooperative Filter Attention): 本文提出基于协同过滤注意力机制的方法。

表 4 中实验结果由两部分组成,其中第一部分表示模型在只考虑评论内容的情況下进行的验证,第二部分表示加入用户信息和产品信息后的实验结果。从表 4 中可以看出:

表 4 实验结果与对比

	IMDB		Yelp2013		Yelp2014	
	Acc	RMSE	Acc	RMSE	Acc	RMSE
Majority	0.196	2.495	0.411	1.060	0.392	1.097
Trigram	0.399	1.783	0.569	0.814	0.577	0.804
TextFeature	0.402	1.793	0.556	0.845	0.572	0.801
AvgWordvec+SVM ^[19]	0.304	1.985	0.526	0.898	0.531	0.893
SSWE+SVM	0.312	1.973	0.549	0.849	0.557	0.851
Paragraph Vector ^[21]	0.314	1.814	0.554	0.832	0.564	0.802
RNTN+Recurrent ^[8]	0.401	1.764	0.574	0.804	0.582	0.821
LSTM	0.434	1.475	0.620	0.697	0.634	0.680
Trigram+UPF ^[20]	0.404	1.764	0.570	0.803	0.576	0.789
TextFeature+UPF	0.402	1.774	0.561	1.822	0.579	0.791
UPNN ^[14]	0.435	1.602	0.596	0.784	0.608	0.764
UPDMN ^[22]	0.465	1.351	0.639	0.662	0.609	0.731
LSTM+CBA ^[23]	0.489	1.365	0.638	0.697	0.641	0.678
LSTM-CFA	0.520	1.282	0.645	0.699	0.670	0.653

(1) 第一部分模型中既有传统的机器学习方法也有神经网络模型。在 Yelp2013 和 Yelp2014 数据集上神经网络效果提升明显,可以发现神经网络与传统机器学习相比可以提取出更多的文本语义信息。第二部分模型中加入用户信息和产品信息后,神经网络模型表现更加出色,在三个数据集上均有

很大的提升,可见神经网络模型在处理时序数据时可以表征出更加丰富的文本信息。

(2) 表 4 中第二部分中 LSTM+CBA 表示使用认知模型作为注意力机制,在 Yelp2014 数据集上提升 3.2% 精准率,可以发现加入模拟人类认知模型对实验影响较大。本文提出在层次 LSTM 模型中加入协同过滤方法(LSTM-CFA),从用户和产品信息中提取出更深层次的用户个性和产品属性等信息。在实验效果中可以发现本文方法在三个数据集上精准率最高提升 3.1%,验证了本文提出方法的有效性。

(3) 表 4 中 LSTM 模型表示没有使用注意力机制的模型。与加入注意力机制模型 LSTM-CFA 模型相比较,在三个数据集上分类性能均有明显提升。其中,在 IMDB 数据集上有 8.6% 提升。可以发现加入用户个性信息和产品属性信息后,模型能够更加准确地提取出文本中的信息,以表征文档特征。这既验证了本文加入注意力机制模型后的有效性,也表明在评论数据集中用户和产品信息的重要性。

本文针对在协同过滤(CF)算法中得到的用户兴趣分布矩阵中的稀疏问题,进行了优化处理。表 5 是使用不同方法优化用户兴趣分布矩阵的结果的对比。其中 CF-DC(direct cut)表示直接对用户兴趣分布矩阵进行截取,得到用户信息矩阵和产品信息矩阵;CF-SVD 表示直接使用 SVD 公式中 Σ 矩阵保留奇异值较大部分后再分别与 U 和 V 矩阵相乘,得到不同维度的用户信息和产品信息;考虑到稀疏矩阵问题,CF-SVDS 表示使用基于 ARPACK 库进行奇异值分解,该算法主要针对大规模稀疏矩阵或者结构化矩阵进行特征值求解;CF-N(normalization)表示在方法 CF-SVDS 基础上对用户打分数据进行归一化处理。由表中数据可以看出,在 Yelp2013 和 Yelp2014 数据集上方法 CF-SVDS 提升较明显,在 IMDB 数据集上分数数据的归一化处理效果较为明显。由此验证了本文中对用户兴趣分布矩阵进行优化方法的有效性。

表 5 不同优化方法对比

	IMDB		Yelp2013		Yelp2014	
	Acc	RMSE	Acc	RMSE	Acc	RMSE
CF-DC	0.505	1.347	0.650	0.689	0.668	0.657
CF-SVD	0.517	1.316	0.634	0.727	0.665	0.669
CF-SVDS	0.511	1.335	0.645	0.694	0.671	0.657
CF-N	0.520	1.282	0.645	0.699	0.670	0.653

4 结论

本文提出了基于协同过滤 Attention 机制的情感分析模型。通过协同过滤中基于物品相似性方法构建用户兴趣分布矩阵作为模型注意力机制,并将经过 SVD 分解优化后的矩阵加入层次 LSTM 模型中。在词语级别和句子级别中分别提取语义信息,以便更好地完成文档级别情感分析任务。根据在三个数据集上实验结果可以发现在神经网络模型中加入协同过滤算法后可以表征更加丰富的文本信息,并且验证了 SVD 分解优化的重要性。之后的工作我们将着力于进一步提升模型的泛化能力。例如,在小样本、样本类别不均衡的特殊情况下,扩展其适用范围。

参考文献

- [1] 赵妍妍,秦兵,刘挺. 文本情感分析[J]. 软件学报, 2010, 21(8):1834-1848.
- [2] Seroussi Y, Zukerman I, Bohnert F. Collaborative Inference of Sentiments from Texts[C]//Proceedings of the International Conference on User Modeling, Adaptation, and Personalization. Springer-Verlag, 2010: 195-206.
- [3] Diao Q, Qiu M, Wu C Y, et al. Jointly modeling aspects, ratings and sentiments for movie recommendation (JMARS)[C]//Proceedings of the ACM Sigkdd International Conference on Knowledge Discovery & Data Mining. ACM, 2014:193-202.
- [4] Hinton G E. Learning distributed representations of concepts[C]//Proceedings of the CogSci. Amherst, MA, 1986: 1-12.
- [5] Mikolov T, Chen K, Corrado G, et al. Efficient Estimation of Word Representations in Vector Space[J]. arXiv preprint arXiv. 2013:1301.3781.
- [6] Socher R, Pennington J, Huang E H, et al. Semi-supervised recursive autoencoders for predicting sentiment distributions[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2011: 27-31.
- [7] Socher R, Huval B, Manning C D, et al. Semantic compositionality through recursive matrix-vector spaces [C]//Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. 2012:1201-1211.
- [8] Socher R, Perelygin A, Wu J Y, et al. Recursive deep models for semantic compositionality over a sentiment

- treebank[C]//Proceedings of Conference on Empirical Methods in Natural Language Processing. 2013:1631-1642.
- [9] Kim Y. Convolutional Neural Networks for Sentence Classification[J]. arXiv preprint arXiv, 2014: 1408. 5882.
- [10] Kalchbrenner N, Grefenstette E, Blunsom P. A Convolutional Neural Network for Modelling Sentences [J]. arXiv preprint arXiv, 2014: 1404. 2188.
- [11] Tai K S, Socher R, Manning C D. Improved Semantic Representations From Tree-Structured Long Short-Term Memory Networks [J]. Computer Science, 2015: 5(1): 36.
- [12] Tang D, Qin B, Liu T. Document Modeling with Gated Recurrent Neural Network for Sentiment Classification[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. 2015: 1422-1432.
- [13] Yang Z, Yang D, Dyer C, et al. Hierarchical Attention Networks for Document Classification[C]//Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2017:1480-1489.
- [14] Wenliang Gao, Naoki Yoshinaga, Nobuhiro Kaji, et al. Modeling user leniency and product popularity for sentiment classification[C]//Proceedings of the IJCNLP, 2013: 1107-1111.
- [15] Zhou C, Sun C, Liu Z, et al. A C-LSTM Neural Network for Text Classification[J]. Computer Science, 2015:1(4):39-44.
- [16] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural Computation, 1997: 9(8): 1735-1780.
- [17] 鲁 骁,王书鑫,王 斌,鲁 凯. 一种融合地理位置信息的协同过滤推荐算法[J]. 中文信息学报, 2016, 30(2): 64-73.
- [18] 张时俊,王永恒. 基于矩阵分解的个性化推荐系统研究[J]. 中文信息学报, 2017, 31(3): 134-139.
- [19] 来斯惟. 基于神经网络的词和文档语义向量表示方法研究[D]. 北京: 中国科学院大学博士学位论文, 2016.
- [20] Le Q V, Mikolov T. Distributed Representations of Sentences and Documents[C]//Proceedings of the International Conference on Machine Learning, 2014, 4 (Vol. II): 1188-1196.
- [21] Mikolov T, Sutskever I, Chen K, et al. Distributed Representations of Words and Phrases and their Compositionality [J]. Advances in Neural Information Processing Systems, 2013: 26:3111-3119.
- [22] Tang D, Qin B, Liu T. Learning Semantic Representations of Users and Products for Document Level Sentiment Classification [C]//Proceedings of the Meeting of the Association for Computational Linguistics and the International Joint Conference on Natural Language Processing. 2015:1014-1023.
- [23] Dou Z, et al. Capturing User and Product Information for Document Level Sentiment Analysis with Deep Memory Network[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. 2017.
- [24] Long Y, Lu Q, Xiang R, et al. A Cognition Based Attention Model for Sentiment Analysis[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. 2017.



赵冬梅(1991—), 硕士, 主要研究领域为文本情感分析、情感计算。

E-main: dmzhao_phy@foxmail.com



陶建华(1972—), 博士, 研究员, 主要研究领域为人机交互、情感计算。

E-main: jhtao@nlpr.ia.ac.cn



李雅(1984—), 博士, 副研究员, 主要研究领域为人机交互、情感计算。

E-main: yli@nlpr.ia.ac.cn